Name, affiliation, and emails of primary contact person

Qi Chen, Victoria University of Wellington, Qi.Chen@ecs.vuw.ac.nz

Name, affiliation, and emails of additional tutorial organisers

Bing Xue, Victoria University of Wellington, Bing.Xue@ecs.vuw.ac.nz
Mengjie Zhang, Victoria University of Wellington, Mengjie.Zhang@ecs.vuw.ac.nz

Tutorial title: Principle and Applications of Semantic Genetic Programming

Tutorial abstract

Semantic genetic programming is a rapidly growing research track of Genetic Programming (GP). Semantic GP incorporates semantic awareness into GP and explicitly uses more information on the behaviour of programs in the search. When evaluating a program, semantic GP characterises it with a vector of outputs instead of a single scalar fitness value. Research has demonstrated the successfulness of additional behavioural information to facilitate the design of a more effective GP search. In addition, the geometric properties of the semantic space lead to more attractive search operators with better theoretical characteristics. With the geometric information of semantics, the GP dynamics are easier to understand and interpret. Inappropriate behaviours are easier to prevent. All these contribute to making GP a more informed and intelligent method. This tutorial will give a comprehensive overview of semantic GP methods. We will review various ways of integrating semantic awareness in the evolutionary process of GP. In particular, we will introduce geometric semantic GP and review its formal geometric semantic framework, and analyse the theoretical properties of the fitness landscape under this framework. This will be followed by a review of many novel developments of provably good semantic genetic operators. Another aspect is the efficient implementation of semantic search operators, which is still challenging. We will illustrate efficient and concise implementations of these operators. Another focus of this tutorial is to stimulate the audience by showing some promising applicative results that have been obtained so far in many applications of semantic GP including many symbolic regression and classification tasks in the areas of healthcare, civil engineering, natural language processing and so on. We will also identify and discuss current challenges and promising future directions in semantic GP with the hope of motivating new and stimulating contributions.

Tutorial keywords: Evolutionary Computation; Genetic Programming; Semantics; and Geometric Semantic.

Topic overview

Semantic Genetic Programming is a relatively new but rapidly growing research track of Genetic Programming (GP). As an evolutionary computation method, GP performs search and optimisation analogising natural evolution to produce programs to achieve the desired state. Program semantics, which provides grounding for the syntax of a program, is an important concept in computer programming. However, most GP methods manipulate programs only with syntax in mind and ignore the knowledge of their semantics. Recent works have shown that semantic GP which makes use of the semantics of GP programs induces a more aware version of GP.

Semantics plays a crucial role in driving the evolutionary search, and thus semantic GP ends up being a more informed and intelligent method. With semantic awareness, GP dynamics are easier to understand and interpret, and inappropriate behaviours are easier to prevent or correct. This exactly contributes to a very important aspect of Explainable AI, which is a current hot research topic in the machine learning community.

Among these semantic GP methods, a relatively new variant is geometric semantic GP (GSGP) which presents a formal geometric framework for program semantics in different problem domains. The geometric framework also provides the basis for designing provably good semantic genetic operators for GP. GSGP aims to search the semantic space directly. The novel geometric semantic crossover and mutation operators act on the syntax of GP programs to produce offspring with desired semantic properties. The most obvious advantage of GSGP is on producing offspring that will not be worse than their worst parent. But more importantly, it induces a unimodal fitness landscape that facilitates the evolvability of GP. The unimodal fitness landscape with no local optimal works for any problems domains, which is a rare benefit in machine learning methods. Some deep investigations on the usefulness and properties of GSGP have demonstrated its advantage in complex real-world applications. A comprehensive review and introduction of these works will help to stimulate more research interests in this promising research track.

Outline of the tutorial structure, with time allocation and distribution of work

The **2 hours** tutorial will consist of the following parts:

1. **An introduction of semantic genetic programming (GP) [25 mins]**
   - introduce the basic idea and components of GP
   - introduce the basic concepts, history and overview of semantic GP techniques

2. **Indirect Semantic GP methods [30 mins]**
   - semantic awareness in initialisation
   - semantic selection
   - genetic operators implicitly utilise semantics
   - what semantics brings

Break [5 mins]

3. **Direct Semantic Method-Geometric Semantic GP and its applications [45 Mins]**
   - introduce geometric semantic framework including the theoretical basis for developing new geometric semantic operators, the uniform fitness landscape
   - introduce the semantic space, and search directly in the semantic space
   - review various novel geometric semantic operators and the implementations
   - real-world applications of geometric semantic GP including for symbolic regression and classification tasks in the application areas of healthcare, civil engineering and natural language processing.

4. **Challenge and Future Directions [15 Mins]**

Short Bio

Qi Chen is currently a Senior Lecturer in the Evolutionary Computation Research Group, Center for Data Science and AI, School of Engineering and Computer Science at Victoria University of Wellington. Her research focuses on data mining, machine learning, evolutionary computation, symbolic regression, feature manipulation. She has over 70 papers published in fully referred international journals and conferences and most of them are on symbolic modeling. Dr Chen has been serving as a program committee member of over ten international conferences including AAAI, IEEE CEC, IEEE SSCI, Australian AI and SEAL. She is serving as a reviewer of over ten international journals including IEEE Transactions on Cybernetics and IEEE Transactions on Evolutionary Computation.

Bing Xue is a Fellow of IEEE, and currently a Professor in Artificial Intelligence and Program Director of Science in Center for Data Science and AI, the School of Engineering and Computer Science at Victoria University of Wellington. She has over 300 papers published in fully refereed international journals and conferences. She is currently the Chair of IEEE CIS Task Force on Transfer Learning & Transfer Optimization. She is an Associate Editor of several international journals, including IEEE TEVC. Prof Xue organised many special sessions and symposiums in international conferences such as IEEE WCCI/CEC, IEEE SSCI, and ACM GECCO. She has been a chair for many international conferences including program chair for SoCPaR2015 and Australasian AI 2018, finance chair for IEEE CEC 2019, general co-chair for IVCNZ 2020, workshop co-chair for IEEE ICDM 2021, and tutorial co-chair for WCCI 2022, IEEE CEC 2024 Conference Chair.

Mengjie Zhang is a Fellow of Royal Society of New Zealand, a Fellow of IEEE, and currently Professor of Computer Science in Center for Data Science and AI, at Victoria University of Wellington, where he heads the interdisciplinary Evolutionary Computation Research Group. He has been serving as an associated editor or editorial board member for over 10 international journals including IEEE Transactions on Evolutionary Computation, and IEEE Transactions on Cybernetics. He is the Tutorial Chair for GECCO 2024, an AIS-BIO Track Chair for GECCO 2016, an EML Track Chair for GECCO 2017, and a GP Track Chair for GECCO 2020. Since 2012, he has been co-chairing several parts of IEEE CEC, SSCI, and EvoIASP/EvoApplications conference (he has been involving major EC conferences such as GECCO, CEC, EvoStar, SEAL). He has been co-organising and co-chairing many special sessions, and also delivered a keynote/plenary talk for IEEE WCCI 2024, IEEE ICAVSS 2018, DOCSA 2019, IES 2017 and Chinese National Conference on AI in Law 2017.